

Het Nieuwe  
Instituut

A preservation policy for  
the AutoCAD DWG/DXF file format

---

**./DATABLE**

Title	A preservation policy for the AutoCAD DWG/DXF file format
Author(s)	Henk Vanstappen, DATABLE
Date	2019-12-12 v1.0

## Contents

1. Introduction	4
2. Autodesk DWG/DXF file format	5
2.1. AutoCAD model	5
2.1.1. Model kernel	5
2.1.2. External references	6
2.1.3. Metadata	6
2.1.4. Versions and compatibility	7
2.2. DWG file format	7
2.2.1. File format specification	7
2.2.2. File Structure	8
2.2.3. Identification	8
2.2.4. Internal validation of a CAD format	8
2.3. DXF file format	9
2.3.1. File structure	9
2.3.2. Identification	10
2.4. CAD Software for DWG/DXF	10
2.4.1. AutoCAD	10
2.4.2. AutoCAD export formats	11
2.4.3. Proprietary applications for DWG/DXF	12
2.4.4. Open source applications for DWG/DXF	12
2.5. Software libraries for DWG/DXF	12
2.5.1. RealDWG	13
2.5.2. DWG Direct	13
2.5.3. Open Source libraries	13
2.6. DWG/DXF Viewers and converters	13
2.7. Preservation risks	14
3. Preservation strategies for AutoCAD DWG/DXF	16
4. Conclusion: towards a file format policy for DWG	18
4.1. Pre-ingest	18
4.1.1. Technology preservation	18
4.1.2. Dealing with xrefs	18
4.1.3. Normalization scenario's	19
4.1.4. Normalization tooling	19
4.1.5. Normalization validation	20
4.2. Ingest	20

---

4.2.1. Format identification	20
4.2.2. Format validation	21
4.2.3. Metadata extraction	21
4.3. Preservation planning	21
4.4. Access	22
4.4.1. Create DIP's	22
4.4.2. Emulation	22
4.5. Technology watch	23
5. Resources	24
6. Addenda	27
6.1. AutoCAD software history	27
6.2. File format history	28
6.3. DWG magic numbers	29
6.4. DWG and DXF file format specifications in PRONOM	30

## 1. Introduction

Het Nieuwe Instituut (The New Institute) in Rotterdam is a museum for architecture, design and digital culture, and a platform for the creative industry. Exhibitions and debates are organized on the various design disciplines such as graphic design, product design, games, fashion, (interior) architecture, urban design and landscape architecture. Het Nieuwe Instituut was created on January 1, 2013 from a merger of the Netherlands Architecture Institute (NAi), Premsela Institute for design and fashion and the Virtual Platform, knowledge institute for e-culture.<sup>1</sup>

Het Nieuwe Instituut aims to acquire, select, preserve and make accessible archives in a reasoned manner within its collection policy, acquisition policy and Preservation policy. In 2019, Het Nieuwe Instituut started testing and developing a digital archive facility. The first tests for setting up a file format policy also take place within this framework. Ultimately, Het Nieuwe Instituut must have a strategy for functional preservation (migration, emulation). In this context, choices will have to be made within the framework of the Preservation Policy with regard to the preservation of essential characteristics of the information objects.

Practical experience and previous analysis showed that AutoCAD Drawing (DWG) is used very frequently within the architectural design process and as an exchange format between different CAD applications. Due to the frequent use and the high heritage value, Het Nieuwe Instituut wants to develop a policy with regard to this format. Based on the acquired knowledge about the DWG file format and the AutoCAD software, Het Nieuwe Instituut wants to:

- (1) have a study carried out into the sustainability risks associated with files in the DWG file format;
- (2) have an analysis carried out with regard to consequences for the preservation strategy that Het Nieuwe Instituut develops;
- (3) obtain conclusions and advice on a migration strategy to be established towards sustainable file formats for DWG files.

Sustainability risks of CAD files in general have already been described in previous studies. Concrete guidelines (file format policies) that take into account the requirements of the archiving organization - such as determining the significant properties and the consequences this has with regard to the migration strategy, however, were lacking.

This study attempts to fulfill this need and - based on research of the DWG/DXF file format - gives a number of components of a file format policy for the AutoCAD file format, taking into account the global preservation strategy of Het Nieuwe Instituut and the present or existing infrastructure to be developed.

---

<sup>1</sup> <http://hetnieuweinstituut.nl/>

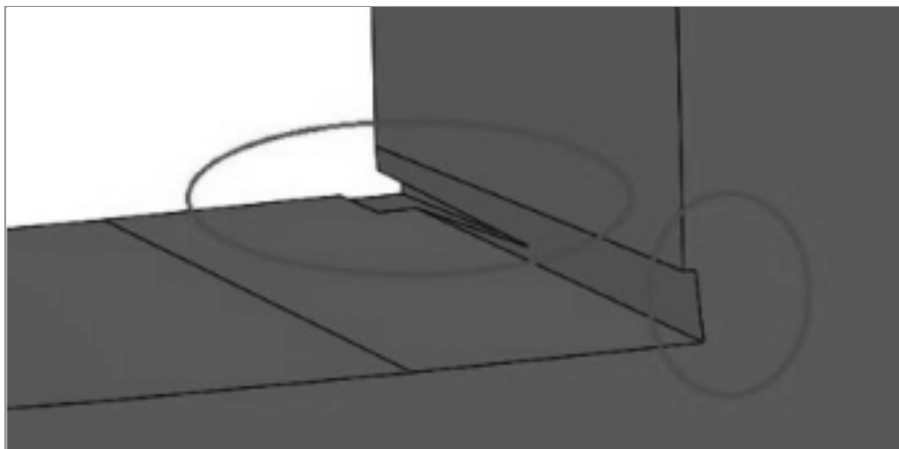
## 2. Autodesk DWG/DXF file format

### 2.1. AutoCAD model

The AutoCAD model, as it is defined in a DWG file and its dependencies, holds all the significant properties that can be selected for retention in function of the designated community.

#### 2.1.1. Model kernel

CAD software uses highly complex mathematical techniques to define and render three dimensional models and their properties. Examples of these techniques are B-spline or NURBS equations, non-parametric equations, or combinations of both. As Smith (2009) puts it, CAD files do not describe a shape as such, but give a recipe of how the shape should be built. The way in which the CAD software can describe a shape is largely determined by the software kernel that is used. Translating 3D models into a format that depends on another kernel therefore runs a greater risk of errors. AutoCAD is based on the ShapeManager kernel, which was forked from the ACIS kernel version 7 in 2001 (Strong, 2019).<sup>2</sup> Strong (2019) therefore states as a rule of thumb that conversions to another format are less likely to cause errors when the target format is based on the same kernel. Errors (or artifacts) that can occur when transforming to another kernel are sliver edges, zero-area faces, duplicate vertices and so on (illustration).



**Sliver faces after conversion<sup>3</sup>**

The AutoCAD kernel supports Brep or border representation. Brep concerns the boundary between fixed and non-fixed geometry, where the fixed geometry is a set of interconnected surfaces. This is a mathematically precise representation of geometry. The opposite of Brep is sometimes referred to as Vrep (visual representation), which offers only an approximation of geometry. AutocAD DWG also supports Vrep.<sup>4</sup>

---

<sup>2</sup> Other common formats are Parasolid (e.g. MicroStation), SMLib of CGM.

<sup>3</sup> Image from <https://www.engineersrule.com/advanced-breakdown-features-solidworks-fillet-tool/>

<sup>4</sup> Vrep formats include Obj, STL, 3D XML, 3D PDF, Collada (.dae) and PLY.

### 2.1.2. External references

In AutoCAD a user can insert any drawing file as an external reference or xref in the current drawing. With xrefs, changes made in the referenced drawing are reflected in the current drawing. Attached xrefs are linked to, but not actually inserted in, another drawing. Any changes to a referenced drawing are displayed in the current drawing when it is opened or reloaded. A drawing file can be attached as an xref to multiple drawings at the same time. Conversely, multiple drawings can be attached as referenced drawings to a single drawing (Autodesk Knowledge Network, 2019).

In many cases, a drawing in the .DWG format will be added as a reference. However, it is also possible to use other file formats:

- Image: Images such as: .BMP, .JPG, .PNG, .TIFF, etc.
- DWF: the Design Web Format: .DWF, .DWFX.
- DGN: the Microstation file format.
- PDF: the Adobe Portable Format document format.
- Point Cloud: .RCP and .RCS files.
- Coordination Model: .NWD and .NWC files from Navisworks.

Xrefs are identified with absolute or relative paths. An absolute path is a fully specified hierarchy of folders that locates the external reference. An absolute path includes the local hard drive letter or the network server drive letter. Relative paths are partially specified folder paths that assume the current drive letter or the folder of the host drawing. This is the most flexible option, and enables you to move a set of drawings from your current drive to a different drive that uses the same folder structure.

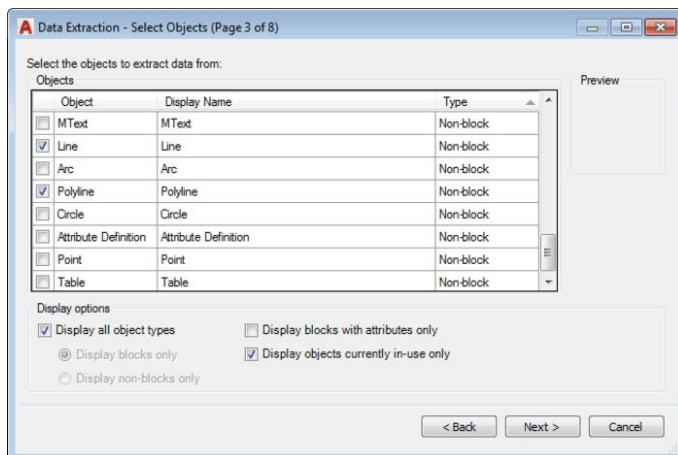
Xrefs can be embedded in a DWG file as separate layers with the binding function.

A Block is a separate drawing inserted into the current drawing as a complete entity. Changes made to the original block will not translate to current drawing.

### 2.1.3. Metadata

Metadata (also called Product Manufacturing Information or PMI) refers to data such as Geometric dimensioning and tolerancing (GD&T), dimensions, and notes which are attached to the solid model (Strong, 2019).

Data extraction is the ability to extract data from objects in your drawing or multiple drawings. AutoCAD provides a Data Extraction Wizard that controls the extraction of that data (illustration).



### AutoCAD data extraction wizard

#### 2.1.4. Versions and compatibility

Like most software, new versions of AutoCAD are regularly released, which also support functionalities. To support these changes, updates of the corresponding file format are also made on a regular basis. Since AutoCAD version 2000, Autodesk changes its DWG file format after three versions. However, from 2013, Autodesk decided to keep the same DWG format (AC1027) for 5 versions, until AutoCAD 2018 (AC1032).

AutoCAD is backwards compatible, meaning files created in any release can be opened and edited in the same or any later release. The software is not forward compatible, but AutoCAD supports the conversion to older format versions.

Binary DXF files can be read only by AutoCAD Release 10 or more recent versions.

## 2.2. DWG file format

DWG is the AutoCAD file format and is used internally by Autodesk in AutoCAD, Revit, Inventor etc. as well as many third party applications (Sheikh, 2019). The DWG file format has evolved with the time since its formal introduction in 1982 (CADAZZ, 2014). Autodesk licensed the DWG file format, which was developed by Mike Riddle in 1970, as the basis for AutoCAD. In 1994 Autodesk introduced 3D solid modelling options simultaneously with the emergence of Windows NT, which made the use of powerful applications more accessible.

### 2.2.1. File format specification

The official DWG specification is undisclosed and proprietary. The Open Design Alliance (ODA, a grouping of a number of Autodesk competitors) therefore decided to reverse engineer the DWG file format (Day, 2006).

The Open Design Specification for DWG files serves AutoCAD's undocumented and proprietary DWG file format. The specification includes DWG file format versions 13 up to and including version 2013. The specification is able to read and write .dwg files, but has some limitations., e.g. the content of the 53 bytes section before the second header is yet unknown (Open Design Alliance, 2018).

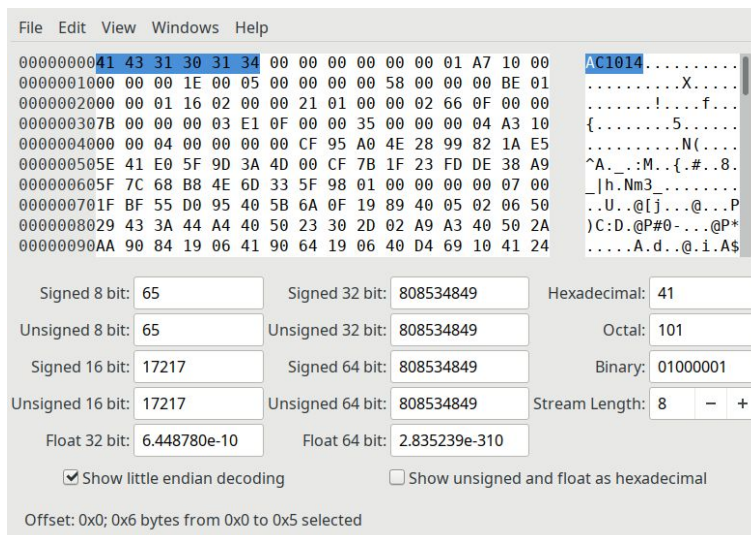
### 2.2.2. File Structure

DWG files usually include information about the image coordinates and any metadata associated with it. The file structure of DWG file format is summarized as follows (Sheek, 2019; Open Design Alliance, 2018).

- Header: the file header consists of DWG Header variables (including the format and version statement) and information about Cyclic Redundancy Check (CRC) which is used for the error detection.
- Class Definitions: information such as class metadata size of class data area, class number and checksums.
- Image Data: the metadata for this section depends on the specific .dwg type.
- Object Data: the object data consists of a complete list of table entities, dictionary entries, etc. corresponding to the existing list of objects.
- Object Map: location of each object in the file is specified in this section of file.
- Second Header: a duplicate of the file header section towards the end of the DWG file.

### 2.2.3. Identification

AutoCAD files can be identified by its extension (.dwg or .dxf). A more reliable and specific identification method is based on the file signature or magic number located in the header of the file. Magic numbers can be used to identify the format, as well as the software version. The DWG and DXF file formats have entries in the National Archives PRONOM registry.<sup>5</sup>



**File opened in HEX viewer with visible magic number**

### 2.2.4. Internal validation of a CAD format

CRC is a mechanism to control the integrity of bitstream (e.g. a file). CRC's are added to the header of a DWG file to support a software function that controls the integrity of different sections of the file.

<sup>5</sup> <http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>



With the release of AutoCAD R14.01 in 1998, Autodesk added another file verification through a function called DWGCHECK that embedded an encrypted checksum and product code, called a WaterMark by Autodesk, into DWG files created by the program. But as a file created with another software library may very well be valid, this function is not very informative.

In 2006 Autodesk modified AutoCAD 2007, to include 'TrustedDWG technology' to embed the text string "Autodesk DWG. This file is a Trusted DWG last saved by an Autodesk application or Autodesk licensed application" into the DWG files. The purpose of this was to help Autodesk software users ensure that these files were created by an Autodesk or RealDWG application, which should help in reducing the risk of incompatibilities (Sheikh, 2019).

## 2.3. DXF file format

DXF was originally introduced in December 1982 as part of AutoCAD 1.0, and was intended to provide an exact ASCII based representation of the data in the AutoCAD native DWG file format. As AutoCAD has become more complex, certain object types, including ACIS solids and regions, are not supported (AutoCAD DXF, n.d.) As all information is encoded in ASCII, a DXF file usually requires more storage space than a DWG file.

The specification of the DXF file format is available from Autodesk (Autodesk, 2011). The latest available version dates from 2011.

DXF files can be either ASCII or binary format. Unlike ASCII DXF files, which entail a trade-off between size and floating-point accuracy, binary DXF files preserve the accuracy in the drawing database. Binary DXF files are reported to be about 25% more compact. (Autodesk, 2011).

### 2.3.1. File structure

Essentially, a DXF file is composed of pairs of codes and associated values. The codes, known as group codes, indicate the type of value that follows. Using these group code and value pairs, a DXF file is organized into sections composed of records, which are composed of a group code and a data item. Each group code and value are on their own line in the DXF file (Autodesk, 2011).

A DXF file (version 2011) is composed of the following sections :

- HEADER section. Contains general information about the drawing. It consists of an AutoCAD database version number and a number of system variables. Each parameter contains a variable name and its associated value.
- CLASSES section. Holds the information for application-defined classes, whose instances appear in the BLOCKS, ENTITIES, and OBJECTS sections of the database. A class definition is permanently fixed in class hierarchy.
- TABLES section contains definitions for the following symbol tables:
  - APPID (application identification table)
  - BLOCK\_RECORD (block reference table)
  - DIMSTYLE (dimension style table)
  - LAYER (layer table)
  - LTYPE (linetype table)
  - STYLE (text style table)
  - UCS (user coordinate system table)

- VIEW (view table)
- VPORT (viewport configuration table)
- BLOCKS section. Contains block definition and drawing entities that make up each block reference in the drawing.
- ENTITIES section. Contains the graphical objects (entities) in the drawing, including block references (insert entities).
- OBJECTS section. Contains the non graphical objects in the drawing. Objects are similar to entities, except that they have no graphical or geometric meaning. Examples of entries in the OBJECTS section are dictionaries that contain mline styles and groups.
- THUMBNAILIMAGE section. Contains the preview image data for the drawing. This section is optional.

### 2.3.2. Identification

The identification of the ASCII DXF file format and version is documented in the \$ACADVER variable in the header section (Autodesk 2011):

- AC1006 = R10;
- AC1009 = R11 and R12;
- AC1012 = R13;
- AC1014 = R14;
- AC1015 = AutoCAD 2000;
- AC1018 = AutoCAD 2004;
- AC1021 = AutoCAD 2007;
- AC1024 = AutoCAD 2010

A binary DXF file begins with a 22-byte sentinel consisting of the following string:

AutoCAD Binary DXF

## 2.4. CAD Software for DWG/DXF

### 2.4.1. AutoCAD

AutoCAD is a commercial computer-aided design (CAD) and drafting software application. Developed and marketed by Autodesk, AutoCAD was first released in December 1989 as a desktop application. AutoCAD is available in two versions: the full fledged AutoCAD and AutoCAD LT. The latter does not support 3D. AutoCAD is available in a number of variants (aka toolsets or verticals): AutoCAD Architecture, AutoCAD Electrical, AutoCAD Map 3D, AutoCAD Mechanical, AutoCAD MEP, AutoCAD Plant 3D en AutoCAD Raster Design.

Autodesk offers two types of access, single-user and multi-user, with their own associated license type. Users with stand-alone licenses must connect to the internet every 30 days to validate their Autodesk ID, but with the exception of cloud-based services, the software works offline for up to 30 days.

The license agreement allows the user to make one archival copy of the solely for backup and archival purposes and solely for the duration of a subscription. The structure and organization, the underlying algorithms and other internals, the protocols, data structures and other externals, and the source code of the Offerings and the APIs constitute proprietary and confidential information of

Autodesk. A user must agree not to engage in any decompiling, disassembling or other reverse engineering or otherwise attempting to discover, learn or study the structure or organization, underlying algorithms or other internals, the protocols, data structures or other externals, or the source code of the application.

#### 2.4.2. AutoCAD export formats

AutoCAD natively supports the following export formats:

Format	Description	Notes
3D DWF (*.dwf) 3D DWFx (*.dwfx)	Autodesk Design Web Format	3DDWF is a compressed, proprietary file format for viewing and inspecting CAD files.
ACIS (*.sat)	ACIS solid object	Exports trimmed NURBS surfaces, regions, and 3D solids to an ACIS file in ASCII (SAT) format. Other objects, such as lines and arcs, are ignored.
DXX Extract (*.dxx)	Attribute extract DXF™	Extracts attribute information from a drawing and creates a separate text file for use with database software.
Encapsulated PS (*.eps)	Encapsulated PostScript	Export of a file in PostScript format as an EPS file, handles some objects specially, e.g. a 2D (planar) polyline with uniform width is output as a PostScript stroked path. The PostScript end cap and miter limit variables are set to approximate the segment joining.
IFC (*.ifc)	Industry Foundation Classes	The core AutoCAD program is not able to export to this file format. AutoCAD Architecture, AutoCAD MEP or Autodesk Civil 3D have an IFC Export feature built in.
IGES (*.iges; *.igs)	IGES	Under ideal conditions, translation preserves the appearance and functionality of entities. However, this process has limitations, and some data will not be preserved in a round trip to IGES and back. For example, when exporting to IGES, a 2D polyline is translated as IGES entity 106:12. When importing from IGES, entity 106:12 translates to a spline. Hence, the resulting drawing may not be identical to the original drawing.
Lithography (*.stl)	Solid object stereolithography	The 3D solid data is translated to a faceted mesh representation consisting of a set of triangles and saved to an STL file.
STEP		The core AutoCAD program is not able to export to this file format. AutoCAD Mechanical toolset exports to STEP versions AP214 and AP203E2
V7 DGN (*.dgn)	MicroStation DGN	The export process translates basic DWG file

		data into the corresponding DGN file data, and specialized data as a best fit. There are several translation options to determine how data is translated during the export process.
V8 DGN (*.dgn)	MicroStation DGN	see above

**Notes:**

- (1) AutoCAD also supports the export of raster image file formats (TIFF, PNG, JPEG, BMP) and PDF.
- (2) Export to ACIS, DXX, IGES and STL is not supported in AutoCAD LT.
- (3) Export to other formats may be available through third party components, available in the Autodesk App Store.<sup>6</sup>

**2.4.3. Proprietary applications for DWG/DXF**

There is a multitude of applications that support importing, editing and exporting of DWG/DXF format. Some of them are based on the Autodesk software library, while others use the ODA library or an open source variant. It is often unclear which engine is used. Below we list a few:

- ABViewer by Cadsofttools
- Alibre Design by Alibre, LLC
- AllyCAD by Knowledge Base
- ArchiCAD by Graphisoft
- BricsCAD by Bricsys
- IntelliCAD by IntelliCAD Technology Consortium
- MicroStation by Bentley Systems
- Autodesk Revit
- Rhinoceros 3D by Robert McNeel and Associates
- SketchUp by Trimble
- SolidWorks by SolidWorks Corp.
- Solid Edge by Siemens PLM Software

**2.4.4. Open source applications for DWG/DXF**

FreeCAD is a free and open-source application that can work with the DXF. FreeCAD's support for the DWG file format has been problematic due to software license compatibility problems with the GNU LibreDWG library (FreeCAD, n.d.)

LibreCAD is a free and open-source (GPLv2) 2D CAD application that can open DWG and DXF files. LibreCAD uses the libdxfw software library.<sup>7</sup>

**2.5. Software libraries for DWG/DXF**

A software library is a suite of data and programming code that is used to develop software programs and applications. It is designed to assist both the programmer and the programming language

<sup>6</sup> [https://apps.autodesk.com/ACD/en/List/Search?facet=\\_\\_category%3a%3aTranslator](https://apps.autodesk.com/ACD/en/List/Search?facet=__category%3a%3aTranslator)

<sup>7</sup> <https://wiki.librecad.org/>

compiler in building and executing software.<sup>8</sup> Autodesk as well as its competitors have developed software libraries that can read, process and export DWG and DXF files.

### 2.5.1. RealDWG

The RealDWG developer toolkit is a software library that allows C++ and .NET developers to read and write AutoCAD software DWG and DXF files. RealDWG contains the APIs for reading and writing AutoCAD DWG and DXF files. RealDWG does not contain support for viewing and access to the AutoCAD user interface. RealDWG is used to create host applications and does not require the presence of AutoCAD software.

The most current version of RealDWG provides compatibility with AutoCAD DWG files, including read and write support for AutoCAD releases since AutoCAD Release 14 and drawing enhancements available with the most current version of AutoCAD.

### 2.5.2. DWG Direct

The Open Design Alliance's Drawings SDK is a development toolkit that provides access to all data in DWG through an object-oriented API, allows creating and editing of a DWG file. The library is used in products such as Bricsys and Intellicad.<sup>9</sup>

### 2.5.3. Open Source libraries

The open source community has made several efforts to provide a truly open software library that can handle DWG files:

- (1) GNU LibreDWG (forked in late 2009 from libDWG) can read most (!) parts of DWG files from version R13 up to 2004. But as the libreDWG library is released under the GNU GPLv3 it can't be used by most targeted open source software, like FreeCAD, LibreCAD and Blender, due to a GPLv2/GPLv3 license incompatibility. The project has stalled since 2011.<sup>10</sup>
- (2) LibDWG - free access to DWG project was reactivated in september 2013. It was re-forked from LibreDWG. However, since march 2015 no update was released and it got abandoned again.<sup>11</sup>
- (3) A GPLv2 licensed alternative is the libdxfrw project, which can read simple 2D DWG files.<sup>12</sup>

## 2.6. DWG/DXF Viewers and converters

Both Autodesk and ODA have made viewers available that can be downloaded for free. Web based viewers are available as well, e.g. Autodesk Viewer.<sup>13</sup>

- (1) Autodesk DWG Trueview<sup>14</sup> is a freeware stand-alone DWG viewer with DWG TrueConvert software included, built on the same viewing engine as AutoCAD software and capable of viewing and converting between different DWG and DXF versions.

---

<sup>8</sup> <https://www.techopedia.com/definition/3828/software-library>

<sup>9</sup> <https://www.opendesign.com/products/drawings>

<sup>10</sup> <https://www.gnu.org/software/libredwg/>

<sup>11</sup> <https://libdwg.sourceforge.io/en/index.html>

<sup>12</sup> <https://github.com/LibreCAD/libdxfrw>

<sup>13</sup> <https://viewer.autodesk.com/>

<sup>14</sup> <https://www.autodesk.com/products/dwg>

- (2) Autodesk Design Review<sup>15</sup> software adds a possibility to open DWG files in Design Review to take advantage of measure and markup capabilities, sheet set organization, and status tracking.
- (3) ODA Drawings Explorer<sup>16</sup> is a freeware standalone viewer for DWG files. It is intended for rendering and testing and runs on Windows, Linux, macOS.
- (4) ODA File Converter application has a graphical interface and a command-line interface, which is capable of batch converting between different versions of DWG and DXF.<sup>17</sup> If the audit flag is enabled, an audit/repair operation will be applied to each file as it is loaded. It runs on Windows, Linux and macOS.

## 2.7. Preservation risks

The sustainability of a file format has been defined by different authors, with the objective of the selection of archival formats, i.e. file formats that can be accepted in a digital preservation system (Rog & van Wijk, 2008; Todd, 2009; Folk & Barkstrom, 2003).

Based on the analysis of the AutoCAD software and the DWG file format, it is possible to identify the preservation risks related to a file format. To determine these risks we use a list of criteria to determine whether a format is suitable as an archive format:

- (1) Adoption — the extent to which the format is in widespread use, e.g; the availability of software or software libraries;
- (2) Platform independence — the extent to which the format is independent of specific support from hardware and software;
- (3) Disclosure — the extent to which the file format specification is in the public domain;
- (4) Transparency — the readiness with which the file format can be inspected or interrogated to discover its identity and attributes, as against where it is obscured by compression, ‘wrapper’ data architectures or other techniques;
- (5) Metadata support — the extent to which descriptive information is supported in machine readable form within the format. This includes OAIIS representation information and occasionally how far the file format supports the recording of management processes it has been subject to.

Adoption of AutoCAD is high: the AutoCAD software and DWG file format is used by architects and designers all over the world. Enlyft (2019) reports a current market share for Autodesk AutoCAD of 37%. There is a global community of engineers, architects and designers, supported by a range of service providers that provide education and support as well as plugins and extensions on the software.

DWG is also supported in other CAD applications, either as a native format or as a format that can be imported and/or exported.

The file format itself is platform independent in that all DWG or DXF files can be read by the Windows

---

<sup>15</sup> <https://www.autodesk.com/products/design-review/overview>

<sup>16</sup> [https://www.opendesign.com/guestfiles/oda\\_drawings\\_explorer](https://www.opendesign.com/guestfiles/oda_drawings_explorer)

<sup>17</sup> [https://www.opendesign.com/guestfiles/oda\\_file\\_converter](https://www.opendesign.com/guestfiles/oda_file_converter)

and macOS versions of the AutoCAD software (Linux is not supported). Users with stand-alone licenses must connect to the internet every 30 days to validate their Autodesk ID, which means there is also a dependency of Autodesk's user management service.

Internally the AutoCAD application depends on few other technologies, except when files in different file formats are embedded using xrefs. As an alternative to AutoCAD, applications and software libraries are available from third parties, albeit with some limitations regarding the support of DWG entities.

As the official DWG file format specification is not published or available otherwise, disclosure is a preservation risk. Thanks to the efforts of the Open Design Alliance, a specification is available that closely approximates the original. However, the ODA specification is not updated to match the most recent version of AutoCAD.

The specification of the DXF counterpart is published by Autodesk, but this version dates back to 2011 and thus is not compatible with most recent versions of DWG. Conversion from DWG to DXF therefore has a risk of information loss.

Also, the source code of the application itself is undisclosed. The reverse engineering efforts have resulted in unofficial but generally speaking complete alternatives in the form of a file format specification and a competing (proprietary) SDK that can handle DWG and DXF files. The initiatives to provide a truly open source software library or application have not been successful so far.

A DWG file can be identified based on the extension .dwg or (even better) by an analysis of the file header, which will also reveal the software version it was produced with. As the structure of the file is documented but as the file is binary, file contents cannot be evaluated without dedicated software. In this respect, ASCII DXF files are more transparent, but as we have seen the ASCII files don't support all information entities in a DWG file.

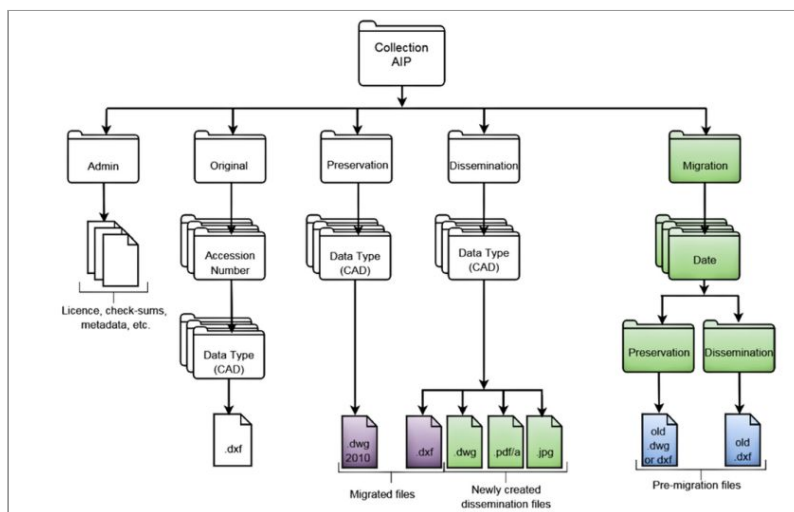
Metadata extraction (except for the file format version, which can be extracted with tools such as DROID or Siegfried) is possible with dedicated software only.



### 3. Preservation strategies for AutoCAD DWG/DXF

Strategies and approaches such as normalization to an archival file format have successfully been implemented for relatively simple file types such as 3D raster images (e.g. TIFF) or text (ODT, XML). Because of its inherent complexity when preserving 3D files, one is confronted with a much higher complexity, stemming from both interdependencies on other systems and translators and complex interrelationships between parts of a single model, e.g. geometry, differences in tolerances, supported entities or metadata, etc. (Chinn, 2009). As attention to the problem of preserving 3D CAD files grew, various studies and projects were conducted that formulated advice on the sustainability of 3D CAD file formats and preservation strategies.

The Archaeology Data Service (ADS) preserves archaeological datasets within the area broadly defined as Archaeology and the Historic Environment. Every file is preserved in a standardised ('normalized') format. The choice of archival formats is based on a mixture of technical considerations, judgements on longevity of format and ease of establishing future migrations. ADS had traditionally used DXF (R14) in AIP and DIP support for textual encoding (ASCII) and its primary purpose as an exchange format which could be used beyond Autodesk software. Due to the fast development of the AutoCAD software, the DXF format has seen almost as many version updates as the proprietary DWG format. As a result, the decision was made in early 2014 to change the ADS archiving policy and adopt DWG version 2010 (AC1024) (Evans, 2016; Green 2016).



**Source: Archaeology Data Service**

Ball (2013) recommends that archives "normalize CAD models to at least one, but ideally two or three, vendor-neutral standard formats," particularly those defined by the STEP (Standard for the Exchange of Product Model Data) international standard (ISO 10303).

The objective of the LOTAR Composites Workgroup is to develop, publish and maintain standards designed to provide the capability to archive and retrieve CAD 3D composite structure in a standard neutral form that can be read and reused throughout the product life cycle, independent of changes



in the IT application environment originally used for creation.<sup>18</sup> The LOTAR standards do not define specific information models for long term preservation of CAD information models. They rely closely on the ISO 10303 STEP Application Protocols. The STEP modular architecture ensures the consistency of the information models subsets common to several ISO 10303 standards.

The first FACADE project (FACADE, 2013) recommends that four versions of CAD files be kept for preservation:

- (1) the original;
- (2) a dissemination format, such as 3D PDF
- (3) a 'heavyweight' standard format, such as IFC or STEP
- (4) a 'lightweight' format, such as IGES, which retains the simple geometry of the model.

The DURAARK project studied the sustainability of the ISO standardized IFC (Industry Foundation Classes) file format for BIMs, which was considered well suited for archival purposes from a sustainability point of view (Lindlar & Saemann, 2014).

Lowet (2016) proposes to save the original (2D) DWG files and to count on the backward compatibility of the DWG for the time being. At the same time, a technology watch must be set up to signal any increased risks (no longer supporting older versions of DWG). In the meantime, the normalization of DWG files can already be prepared. In this scenario, DXF is chosen as the archival format.

Finally, in their representation of managing and providing architecture archives at the Canadian Centre for Architecture (CCA), Stewart & Breitweiser (2019) seem to suggest that no standardization or migration is being performed on DWG and DXF files.

---

<sup>18</sup> <http://www.lotar-international.org>

## 4. Conclusion: towards a file format policy for DWG

In the following chapter we propose a procedure for the handling of AutoCAD DWG and DXF files in a trusted digital archive. The steps of this procedure are loosely based on the OAIS model, as explained in (Walsh, 2015). The steps assume an already defined strategy and overall procedure for management and preservation of digital archives, as defined in Ras, 2018. It also assumes that the organization has an OAIS compliant management system, i.c. Archivemataca.<sup>19</sup> Vital components of a digital archive infrastructure (such as storage media and repository systems) are therefore mentioned only briefly. We highlight those parts that are specific to DWG/DXF files.

Het Nieuwe Instituut does the recording, management, preservation actions and the accessibility actions of digital objects as much as possible automatically. This is feasible with regard to bit preservation. The degree of automation will decrease for research and functional preservation. This generally requires customization.

Although architectural archives with CAD files can differ greatly and require different levels of care and analysis in practice, we try to describe a generic process here, to which exceptions can be imposed depending on specific requirements.

### 4.1. Pre-ingest

In the Pre-ingest phase Het Nieuwe Instituut carries out checks on:

- (1) The technical formats of the information objects supplied;
- (2) The limitations in digital signatures, compression and other technical operations;
- (3) The presence of metadata;
- (4) The agreements with regard to (re) use and access (Ras, 2018).

#### 4.1.1. Technology preservation

Preserving technology approaches attempt to keep data in specific logical or physical formats and use technology originally associated with those formats. This can be achieved by actually preserving the entire environment (hardware, operating system, software, files) to represent the original DWG/DXF models in an archival context (Vanstappen, 2019). This ‘computer museum’ approach is not be considered as a viable preservation strategy. However, it is a good idea to archive software with a view to documenting the functionalities of the application and as a means to perform pre-ingested and ingested processes insofar as these cannot be embedded within the Archivemataca workflows.

#### 4.1.2. Dealing with xrefs

In the pre-ingest phase, files may be taken out of their original environment. When doing so, precautions must be taken to avoid file corruption through the loss of external references (Xrefs). This can be done using the Bind function in AutoCAD, or by changing the xref paths to a relative path and copy both the drawing and its xrefs to a new location.

When an xref is in an other format than DWG/DXF, proper measures must be taken in accordance with the file format policy of the given file format.

---

<sup>19</sup> <https://www.archivemataca.org/>

#### 4.1.3. Normalization scenario's

At the moment, Het Nieuwe Instituut does not impose any restrictions on the number or type of file formats to be included. There is no legal framework for this. Het Nieuwe Instituut will, however, draw up a list of preferred formats.

According to Het Nieuwe Instituut's preservation policy, a conversion from certain formats to a more sustainable format may already follow. Both formats - including the metadata - are saved. At the end of the pre-ingest there is a valid and usable submission information package (SIP).

As described in the overview of existing and recommended preservation strategies, several scenarios with regard to normalization are conceivable:

- (1) No normalization: DWG and DXF files are saved without further intervention. Although DWG does not strictly meet the requirements of an archiving format, this option is certainly acceptable. In its current form and context, DWG has few immediate preservation risks. The chance that the format turns out to be illegible in the short or medium term is very limited. The same applies to DXF, which moreover scores better in terms of transparency and is itself an open standard.
- (2) When the preservation policy requires the normalization of DWG to an open and standardized archiving format, there are several options:
  - (a) The sometimes proposed strategy to convert DWG to ASCII DXF files involves a number of risks. Although DXF is an open specification and is more transparent due to its ASCII coding, the current specification is considerably behind the latest DWG specification. Moreover, the accuracy of ASCII DXF is lower. Loss of information is therefore a real risk. Conversion to DXF is only recommended as an supplementary step, but normalization to another format is preferred.
  - (b) STEP AP203 defines the geometry, topology, and configuration management data of solid models for mechanical parts and assemblies. This file type does not manage Colors and Layers. AP214 has everything a AP203 file includes, but adds colors, layers, geometric dimensioning and tolerance, and design intent. AP214 is considered an extension of AP203.
  - (c) IFC (Industry Foundation Classes) is an open file format for the description of architectural, building and construction industry data. The format is used as the collaboration format in Building information modeling (BIM) projects.

In the present circumstances, however, we estimate the necessity or priority of normalization to be low. Only if support from DWG appears to be uncertainty in the future (cf. technology watch), this normalization step is recommended.

Taking into account the backwards compatibility of DWG, there is currently no need for a migration to a more recent format of DWG.

#### 4.1.4. Normalization tooling

Normalization is Archivematica's primary format preservation strategy. The preservation copies are added to the AIP and the access copies are used to generate a DIP for upload to the access system.

Several different tools are used to complete normalization tasks within Archivematica, depending on the format of the file. None of the built in options are capable of converting DWG/DXF files. If

normalization is required, some alternative options are available to include this step - eventually into Archivematica's automated workflow:

- (1) The ODA File Converter application is capable of DWG/DXF conversion and version migration from the command line.
- (2) DWG TrueView is also capable of executing the same commands, from the (Windows) command line.
- (3) Specialized tools, such as TransMagic<sup>20</sup> should be investigated for normalization to other formats.

#### 4.1.5. Normalization validation

Validation of the result of a normalization process implies that every feature of the original file format (geometry, position, metadata, ...) is checked against the features of the migrated file. This process proves to be very complex, even for a relatively simple format as STEP (Cinn, 2009).

Strong (2018) distinguishes three ways of comparing two models:

- (1) Visual inspection, in which the original is colored green and the revised version is colored red. Both models are switched to wireframe display and visually inspected using brute-force eyeball strength.
- (2) Comparison software built into a CAD application, when available. It is reported to be rudimentary and to lack the precision.
- (3) Use of CAD Comparison software that automatically evaluates the exact degree to which two models have the same geometry, provides a method to authenticate that the two models are the same for all practical purposes and determines how well a grouping of points fits to an existing 3D CAD model.

## 4.2. Ingest

In Het Nieuwe Instituut's preservation policy, at the ingest of the SIP a number of checks and identifications take place (Ras, 2018).

### 4.2.1. Format identification

Identification is the process of analyzing given information about a file to derive its format. In Archivematica, there are three file identification tools supporting two identification methods:

- (1) File extension, a simple script which identifies files by their file extension and thus is not capable of distinguishing different format versions;
- (2) FIDO<sup>21</sup> and Siegfried<sup>22</sup> (default) which identify files by their signature and connect this to a PRONOM ID.

The PRONOM database contains the signatures of the different DWG and DXF versions. Archival software is therefore sufficiently equipped to adequately recognize the formats and versions. However, the most recent versions of DWG are not yet included in the database (the most recent is DWG 2014, cf. Addenda). Het Nieuwe Instituut should take the initiative here to add these signatures to the PRONOM database.

---

<sup>20</sup> <https://transmagic.com/cad-automation/>

<sup>21</sup> <https://github.com/openpreserve/fido/>

<sup>22</sup> <https://www.itforarchivists.com/siegfried>

#### 4.2.2. Format validation

Format validation ensures that files are well-formed and compliant with any relevant format specifications. There are two aspects on the validation process. Firstly a validation can be applied to a file in order to control if it is structured conforming the file format specification. The digital preservation community has produced established tools such as JHOVE and VeraPDF to validate file formats such as WAV, TIFF or PDF.<sup>23,24</sup> For DWG files, validation can be executed with proprietary validators from software producers.

Archivematica contains two validation tools: JHOVE and MediaConch, both of which are not suitable for validating DWG or DXF files. A limited validation can be executed with the normalization tools that are previously mentioned. As mentioned before, AutoCAD actually provides an (overly strict) validation functionality called DWGCHECK.

#### 4.2.3. Metadata extraction

Metadata is information about relationships between and information about information objects. Het Nieuwe Instituut is working on a Metadata Directive. This specifies which metadata must be stored, within which systems and with which mutual relationships (Ras, 2018).

A distinction is made between:

- Descriptive metadata. These describe the information objects on content characteristics;
- Technical metadata. These describe the information objects themselves and are necessary to guarantee long-term access [Characterization];
- Structural metadata. These describe the structure of the archives and mutual relationships within an archive.

Characterization is the process of producing technical metadata for an object. Archivematica's characterization aims both to document the object's significant properties and to extract technical metadata contained within the object. Archivematica has four characterization tools available upon installation. Unfortunately, again none of these tools supports metadata extraction from DWG or DXF files.<sup>25</sup>

Metadata extraction is possible with third party tools (e.g. OpenKM DMS) or the Autodesk metadata viewer.<sup>26</sup> The extent to which these tools can be integrated into the automation workflow of Archivematica must be investigated.

### 4.3. Preservation planning

The e-Depot of Het Nieuwe Instituut will be set up for both bit preservation and functional preservation (Ras, 2018). This is done by:

- (1) Maintaining one original and at least one copy of each bitstream;

---

<sup>23</sup> <https://jhove.openpreservation.org>

<sup>24</sup> <https://verapdf.org>

<sup>25</sup> <https://www.archivematica.org/en/docs/archivematica-1.10/user-manual/preservation/preservation-planning/>

<sup>26</sup> <https://knowledge.autodesk.com/guideref/MAP/2019/learn-explore/GUID-A5F25740-7E04-402E-A5AB-C1177FD8F438>

- (2) Guarantee bitstream integrity (check checksum) and set up a check cycle;
- (3) Being able to prove and document this.

A control mechanism to check the integrity of files can be based on a cryptographic algorithm (checksum) and should be in place at ingest and run at regular intervals to trace any unwanted change of a file (bit rot). The built-in checksum mechanism of DWG does not replace the need for this control mechanism:

- (1) the internal mechanism is only executed when the file is actually opened with an application that supports this mechanism.
- (2) the integrity check must be executed on all manifestations of the document (original file, metadata and - when applicable - the normalized file).

## 4.4. Access

Het Nieuwe Instituut's Preservation policy states that the access functionality supports the accessible, readable and usable offering of information objects. Depending on the designated community or user, the information can be made available in various ways, for example via a viewer or download functionality (Ras, 2018).

### 4.4.1. Create DIP's

When granting access to the DWG/DXF files, the user must be able to visualize and study them. The way in which this is facilitated depends on the requirements of the user and is tailored to the designated community. Access therefore assumes the creation of a DIP, in which the file is presented in a form that may or may not have been modified. Different options are applicable here:

- (1) Creation of a access copy in a common file format, such as PDF or JPEG.
- (2) Representing the original format in a suitable viewer (e.g. Autodesk DWG Trueview or ODA Drawings Explorer).
- (3) Representing the normalized format in a suitable viewer (e.g. Autodesk Viewer<sup>27</sup>).
- (4) Offer the possibility to download the original or normalized file, referring to a suitable viewer.

The use of PDF naturally has certain limitations on the accurate representation of a CAD file - in particular of 3D files. On the other hand, it is an accessible format that is already present on most desktops. The use of a raster image (e.g. JPG) offers an even greater accessibility, but the degree of information loss is even higher. PDF and JPG access files are therefore especially suitable for giving a first impression of the file. Specialized tools are needed to create such access formats (see above under Normalisation).

### 4.4.2. Emulation

An alternative to offering access files or original files in a viewer or as a download is to emulate the entire environment (i.e. operating system and software). This approach offers the advantage that the information object can be displayed in its original context or an approximation thereof. Moreover, this approach does not depend on the availability of viewers. On the other hand, setting up an emulation environment provides additional complexity.

---

<sup>27</sup> <https://viewer.autodesk.com>

The emulation option therefore does not seem to be a priority for the near future. In the longer term operating systems and software may become obsolete. Therefore emulation is a strategy that needs further investigation. This investigation involves not only the technical challenges, but also the legal barriers caused by copyright issues on software and operating systems.

#### **4.5. Technology watch**

Due to the evolution of CAD software, it is important to closely follow technological trends. This Technology watch function includes:

- (1) evolution of the DWG / DXF file format, in particular the backward compatibility of the format and the compatibility with archiving standards such as STEP and IFC;
- (2) following up on developments in software development with regard to CAD applications and viewers;
- (3) following trends in acceptance and market share of the DWG / DXF format;
- (4) following up on innovations in emulation of obsolete software and operating systems;
- (5) monitor the availability of file signatures of new versions of DWG and DXF in external databases (e.g. PRONOM), and possibly take action to add these signatures.

## 5. Resources

- AutoCAD DXF (n.d.). In Wikipedia. Retrieved from [https://en.wikipedia.org/wiki/AutoCAD\\_DXF](https://en.wikipedia.org/wiki/AutoCAD_DXF).
- Autodesk Knowledge Network (2019). About Attaching and Detaching Referenced Drawings (Xrefs). Retrieved from <https://knowledge.autodesk.com/support/autocad/getting-started/caas/CloudHelp/cloudhelp/2016/ENU/AutoCAD-Core/files/GUID-A987D2FF-45BD-474E-99C1-E6316A42F667-htm.html>
- Autodesk (2011). AutoCAD 2012 DXF Reference. Retrieved from [https://images.autodesk.com/adsk/files/autocad\\_2012\\_pdf\\_dxf-reference\\_enu.pdf](https://images.autodesk.com/adsk/files/autocad_2012_pdf_dxf-reference_enu.pdf)
- Ball, Alexander (2013). Preserving Computer-Aided Design (CAD). DPC Technology Watch Report 13-02. Digital Preservation Coalition. Retrieved from <https://www.dpconline.org/docs/technology-watch-reports/896-dpctw13-02-pdf/file>
- CADAZZ (2004). CAD software history. Retrieved from <http://cadazz.com/cad-software-history-1995-1997.htm>
- Chinn, A. (2009). Activities in the Development of Standards and Technology for the Long Term Retention of 3D Data. Retrieved August, 30. Retrieved from <http://www.ukoln.ac.uk/events/ltkr-2007/presentations/a-chinn.pdf>
- Day, Martin (2006). The DWG conundrum. In: AEC Magazine. December(7)
- Enlyft (2019). AutoCAD vs Solidworks: Worldwide Market Share Compared. Retrieved from <https://enlyft.com/autocad-vs-solidworks-worldwide-market-share-compared/>
- Evans, Tim (2016). File obsolescence at the ADS? Retrieved from <https://dpconline.org/docs/miscellaneous/events/2016-events/1546-reformat-timevans/file>
- FACADE (2013) Final Report: FACADE2: MIT and Harvard Collaboration," Harvard Library Lab. Retrieved from [https://osc.hul.harvard.edu/liblab/sites/default/files/325\\_final\\_2013\\_0.pdf](https://osc.hul.harvard.edu/liblab/sites/default/files/325_final_2013_0.pdf)
- Folk, M., & Barkstrom, B. R. (2003, May). Attributes of file formats for long-term preservation of scientific and engineering data in digital libraries. In Joint Conference on Digital Libraries (JCDL), Houston, TX (Vol. 1). Retrieved from [https://www.researchgate.net/publication/228726593\\_Attributes\\_of\\_file\\_formats\\_for\\_long-term\\_preservation\\_of\\_scientific\\_and\\_engineering\\_data\\_in\\_digital\\_libraries](https://www.researchgate.net/publication/228726593_Attributes_of_file_formats_for_long-term_preservation_of_scientific_and_engineering_data_in_digital_libraries)
- FreeCAD (n.d.) In Wikipedia. Retrieved from <https://en.wikipedia.org/wiki/FreeCAD>
- Green, K., Niven, K., & Field, G. (2016). Migrating 2 and 3D Datasets: Preserving AutoCAD at the Archaeology Data Service. ISPRS International Journal of Geo-Information, 5(4), 44. <https://doi.org/10.3390/ijgi5040044>
- Heutelbeck, D., Brunsmann, J., Wilkes, W., & Hundsörfer, A. (2009, June). Motivations and challenges for digital preservation in design and engineering. In First International



- Workshop on Innovation in Digital Preservation (InDP), Austin (Vol. 19). Retrieved from [https://www.researchgate.net/publication/228617821\\_Motivations\\_and\\_challenges\\_for\\_digital\\_preservation\\_in\\_design\\_and\\_engineering](https://www.researchgate.net/publication/228617821_Motivations_and_challenges_for_digital_preservation_in_design_and_engineering)
- Lindlar, M., and Saemann, H. (2014) The DURAARK Project – Long-Term Preservation of Architectural 3D-Data. Annual Conference of the International Committee for Documentation / the International Council of Museums (CIDOC 2014). Dresden, Germany. Retrieved from [http://www.cidoc2014.de/images/sampleddata/cidoc/papers/L-1\\_Lindlar\\_Saemann\\_paper.pdf](http://www.cidoc2014.de/images/sampleddata/cidoc/papers/L-1_Lindlar_Saemann_paper.pdf)
- Lowet, Wim (2016). Bouwstenen voor de archivering van het digitaal archief Maarten van Severen. Vlaams Architectuurinstituut. Retrieved from [https://www.vai.be/volumes/general/mvs\\_20160907\\_rapport\\_v1-3.pdf](https://www.vai.be/volumes/general/mvs_20160907_rapport_v1-3.pdf)
- Open Design Alliance (2018). Open Design Specification for .dwg files. Version 5.4.1. Retrieved from [https://www.opendesign.com/files/guestdownloads/OpenDesign\\_Specification\\_for\\_.dwg\\_files.pdf](https://www.opendesign.com/files/guestdownloads/OpenDesign_Specification_for_.dwg_files.pdf)
- Ras, Marcel (2018). Preservation policy. Rotterdam. Het Nieuwe Instituut.
- Rog, J. and van Wijk, C. (2008). Evaluating file formats for long-term preservation. National Library of the Netherlands; The Hague, The Netherlands.
- Sheikh, Farooq (2019). What is a DWG file? Retrieved from <https://wiki.fileformat.com/cad/dwg/>
- Shubert, H. (2008). Preserving digital archives at the canadian centre for architecture: Greg Lynn's embryological house. Architecture et archives numériques. L'architecture à l'ère du numérique: un enjeu de mémoire.
- Smith, MacKenzie (2009). Curating Architectural 3D CAD Models. In: The International Journal of Digital Curation 1 no. 2. Retrieved from <https://doi.org/10.2218/ijdc.v4i1.81>
- Stewart, K., & Breitwieser, S. (2019). SCOPE: A digital archives access interface. Code4Lib Journal, (43). Retrieved from <https://journal.code4lib.org/articles/14283>
- Strong, Brad (2017). Brep vs Visrep Models. Retrieved from <https://transmagic.com/brep-vs-visrep-models/>
- Strong, Brad (2019). Which Geometric Modeling Kernel? Retrieved from <https://transmagic.com/which-geometric-modeling-kernel/>
- Tatum, L. (2002). Documenting Design: A Survey of State-of-the-Art Practice for Archiving Architectural Records. Art Documentation: Journal of the Art Libraries Society of North America, 21(2), 25-31.
- Todd, Malcolm (2009). File formats for preservation. The National Archives (DPC Technology Watch Report Series 09-02).
- Open design Alliance (2018). Open Design Specification for .dwg files Version 5.4.1. Retrieved from [https://www.opendesign.com/files/guestdownloads/OpenDesign\\_Specification\\_for\\_.dwg\\_files.pdf](https://www.opendesign.com/files/guestdownloads/OpenDesign_Specification_for_.dwg_files.pdf)

Vanstappen, Henk (2017). Pre-ingest born digital archief architect Christian Kieckens - Rapport 2: Procedure Identificatie. Antwerp, DATABLE/VAi. Retrieved from [https://www.projectcest.be/w/images/Db-2\\_Procedure\\_identificatie\\_v1\\_6.pdf](https://www.projectcest.be/w/images/Db-2_Procedure_identificatie_v1_6.pdf)

Vanstappen, Henk (2019). SketchUp in digital archives. Software and file format analysis and exploration of the options for digital preservation. Antwerp, Datable/Vlaams Architectuurinstituut.

Walsh, T. (2015). Preservation and Access of Born-Digital Architectural Design Records in an OAIS-Type Archive. Retrieved from [https://www.researchgate.net/publication/228726593\\_Attributes\\_of\\_file\\_formats\\_for\\_long-term\\_preservation\\_of\\_scientific\\_and\\_engineering\\_data\\_in\\_digital\\_libraries](https://www.researchgate.net/publication/228726593_Attributes_of_file_formats_for_long-term_preservation_of_scientific_and_engineering_data_in_digital_libraries)

## 6. Addenda

### 6.1. AutoCAD software history

#### AutoCAD Release History

AutoCAD 1.0 December 1982 (Release 1)

AutoCAD 1.2 April 1983 (Release 2)

AutoCAD 1.3 August 1983 (Release 3)

AutoCAD 1.4 October 1983 (Release 4)

AutoCAD 2.0 October 1984 (Release 5)

AutoCAD 2.1 May 1985 (Release 6)

AutoCAD 2.5 June 1986 (Release 7)

AutoCAD 2.6 April 1987 (Release 8)

AutoCAD R9 September 1987 codename White Album (Release 9)

AutoCAD R10 October 1988 codename Abbey Road (Release 10)

AutoCAD R11 October 1990 codename Let it Be (Release 11)

AutoCAD R12 June 1992 (Release 12)

AutoCAD R13 November 1994 (Release 13)

AutoCAD R14 February 1997 codename Sedona and Pinetop for 14.01 (Release 14)

AutoCAD 2000 March 1999 codename Tahoe (Release 15)

AutoCAD 2000i July 2000 codename Banff (Release 16)

AutoCAD 2002 June 2001 codename Kirkland (Release 17)

AutoCAD 2004 March 2003 codename Reddeer (Release 18)

AutoCAD 2005 March 2004 codename Neo (Release 19)

AutoCAD 2006 March 2005 codename Rio (Release 20)

AutoCAD 2007 March 2006 codename Postrio (Release 21)

AutoCAD 2008 March 2007 codename Spago (Release 22)

AutoCAD 2009 March 2008 codename Raptor (Release 23)

AutoCAD 2010 March 2009 codename Gator (Release 24)

AutoCAD 2011 March 2010 codename Hammer (Release 25)

AutoCAD 2012 March 2011 codename Ironman (Release 26)

AutoCAD 2013 March 2012 codename Jaws (Release 27)

AutoCAD 2014 March 2013 codename Keystone (Release 28)

AutoCAD 2015 March 2014 codename Longbow (Release 29)

AutoCAD 2016 March 2015 codename Maestro (Release 30)

AutoCAD 2017 March 2016 codename Nautilus (Release 31)

AutoCAD 2018 March 2017 codename Omega (Release 32)

AutoCAD 2019 April 2018 codename Pi (Release 33)

### AutoCAD for macOS Releases

AutoCAD for Mac June 1992

AutoCAD for Mac R13 [1994]

AutoCAD 2011 for Mac October 2010 (SledgeHammer)

AutoCAD 2012 for Mac August 2011 (Iron Maiden)

AutoCAD LT 2012 for Mac August 2011 (Ferris)

AutoCAD LT 2013 for Mac August 2012

AutoCAD 2013 for Mac March 2012 (Jaws)

AutoCAD LT 2014 for Mac

AutoCAD 2014 for Mac (Sandstone)

AutoCAD 2015 for Mac (Lightsaber)

AutoCAD 2016 for Mac (Mandalore)

AutoCAD 2017 for Mac (Naboo)

AutoCAD 2018 for Mac Nov 2017

Source: [https://autodesk.blogs.com/between\\_the\\_lines/autocad-release-history.html](https://autodesk.blogs.com/between_the_lines/autocad-release-history.html)

## 6.2. File format history

Version	Internal version name	Software release version name
DWG R1.0	MC0.0	AutoCAD Release 1.0
DWG R1.2	AC1.2	AutoCAD Release 1.2
DWG R1.40	AC1.40	AutoCAD Release 1.40
DWG R2.05	AC1.50	AutoCAD Release 2.05
DWG R2.10	AC2.10	AutoCAD Release 2.10
DWG R2.21	AC2.21	AutoCAD Release 2.21
DWG R2.22	AC1001, AC2.22	AutoCAD Release 2.22
DWG R2.50	AC1002	AutoCAD Release 2.50
DWG R2.60	AC1003	AutoCAD Release 2.60

DWG R9	AC1004	AutoCAD Release 9
DWG R10	AC1006	AutoCAD Release 10
DWG R11/12	AC1009	AutoCAD Release 11, AutoCAD Release 12
DWG R13	AC1012	AutoCAD Release 13
DWG R14	AC1014	AutoCAD Release 14
DWG 2000	AC1015	AutoCAD 2000, AutoCAD 2000i, AutoCAD 2002
DWG 2004	AC1018	AutoCAD 2004, AutoCAD 2005, AutoCAD 2006
DWG 2007	AC1021	AutoCAD 2007, AutoCAD 2008, AutoCAD 2009
DWG 2010	AC1024	AutoCAD 2010, AutoCAD 2011, AutoCAD 2012
DWG 2013	AC1027	AutoCAD 2013, AutoCAD 2014, AutoCAD 2015, AutoCAD 2016, AutoCAD 2017
DWG 2018	AC1032	AutoCAD 2018, AutoCAD 2019, AutoCAD 2020

Source: <https://en.wikipedia.org/wiki/.dwg>

### 6.3. DWG magic numbers

The table below shows the AutoCAD version numbers and their magic number in Hex notation.

Version name	Internal version name	magic number
AutoCAD R1.0 Drawing	MC0.0	4D 43 30 2E 30
AutoCAD R1.2 Drawing	AC1.2	41 43 31 2E 32
AutoCAD R1.40 Drawing	AC1.40	41 43 31 2E 34 30
AutoCAD R2.05 Drawing (new)	AC2.50	41 43 32 2E 35 30
AutoCAD R2.05 Drawing	AC1.50	41 43 31 2E 35 30
AutoCAD R2.10 Drawing	AC2.10	41 43 32 2E 31 30
AutoCAD R2.21 Drawing	AC2.21	41 43 32 2E 32 31
AutoCAD R2.22-20xx Drawing (generic)	AC10	41 43 31 30
AutoCAD R2.22 Drawing (new)	AC1001	41 43 31 30 30 31
AutoCAD R2.22 Drawing (old)	AC2.22	41 43 32 2E 32 32
AutoCAD R2.5 Drawing	AC1002	41 43 31 30 30 32
AutoCAD R2.6 Drawing	AC1003	41 43 31 30 30 33
AutoCAD R9 Drawing	AC1004	41 43 31 30 30 34
AutoCAD R10 Drawing	AC1006	41 43 31 30 30 36
AutoCAD R11-12 Drawing	AC1009	41 43 31 30 30 39
AutoCAD R13 Drawing (subtype 10)	AC1010	41 43 31 30 31 30
AutoCAD R13 Drawing (subtype 11)	AC1011	41 43 31 30 31 31

AutoCAD R13 Drawing	AC1012	41 43 31 30 31 32
AutoCAD R14 Drawing (subtype 13)	AC1013	41 43 31 30 31 33
AutoCAD R14 Drawing	AC1014	41 43 31 30 31 34
AutoCAD 2000-2002 Drawing	AC1015	41 43 31 30 31 35
AutoCAD 2004-2006 Drawing	AC1018	41 43 31 30 31 38
AutoCAD 2007-2009 Drawing	AC1021	41 43 31 30 32 31
AutoCAD 2010-2012 Drawing	AC1024	41 43 31 30 32 34
AutoCAD 2013-2016 Drawing	AC1027	41 43 31 30 32 37
AutoCAD 2018-2019 Drawing	AC1032	41 43 31 30 33 32

Source: <http://mark0.net/soft-trid-e.html>

## 6.4. DWG and DXF file format specifications in PRONOM

File format name	version
AutoCAD Drawing	1
AutoCAD Drawing	1.2
AutoCAD Drawing	1.3
AutoCAD Drawing	1.4
AutoCAD Drawing	2
AutoCAD Drawing	2.1
AutoCAD Drawing	2.2
AutoCAD Drawing	2.5
AutoCAD Drawing	2.6
AutoCAD Drawing	R9
AutoCAD Drawing	R10
AutoCAD Drawing	R11/12
AutoCAD Drawing	R13
AutoCAD Drawing	R14
AutoCAD Drawing	2000-2002
AutoCAD Drawing	2004-2005
AutoCAD Drawing	2010/2011/2012
AutoCAD Drawing	2013/2014
Drawing Interchange File Format (ASCII)	Generic
Drawing Interchange File Format (Binary)	R11/12
Drawing Interchange File Format (Binary)	R13
Drawing Interchange File Format (Binary)	R14

---

Drawing Interchange File Format (Binary)	2000-2002
Drawing Interchange File Format (Binary)	2004-2005
Drawing Interchange File Format (ASCII)	1
Drawing Interchange File Format (ASCII)	1.2
Drawing Interchange File Format (ASCII)	1.3
Drawing Interchange File Format (ASCII)	1.4
Drawing Interchange File Format (ASCII)	2
Drawing Interchange File Format (ASCII)	2.1
Drawing Interchange File Format (ASCII)	2.2
Drawing Interchange File Format (ASCII)	2.5
Drawing Interchange File Format (ASCII)	2.6
Drawing Interchange File Format (ASCII)	R9
Drawing Interchange File Format (ASCII)	R10
Drawing Interchange File Format (ASCII)	R11/12
Drawing Interchange File Format (ASCII)	R13
Drawing Interchange File Format (ASCII)	R14
Drawing Interchange File Format (ASCII)	2000-2002
Drawing Interchange File Format (ASCII)	2004/2005/2006
Drawing Interchange File Format (Binary)	R10
Drawing Interchange File Format (ASCII)	2007/2008/2009
Drawing Interchange File Format (ASCII)	2010/2011/2012
Drawing Interchange File Format (ASCII)	2013/2014
Drawing Interchange Binary Format	1